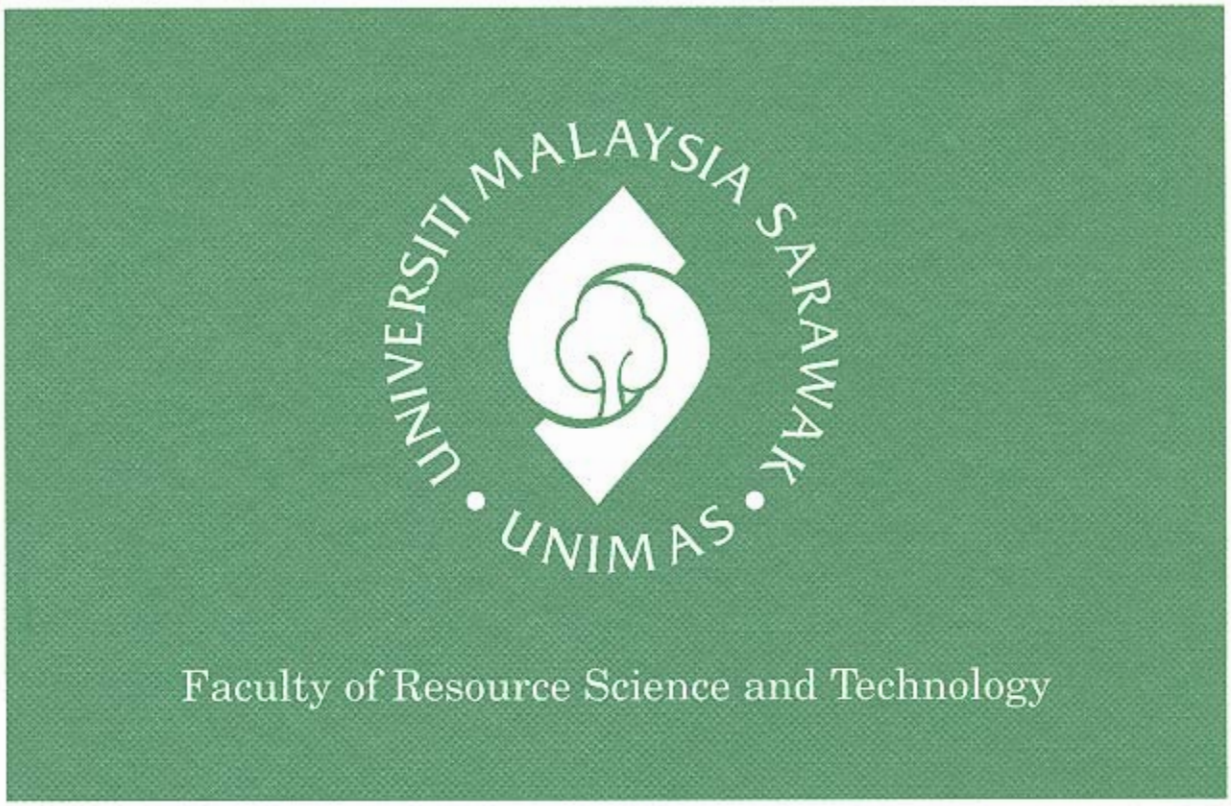

A PHYLOGENETIC STUDY OF THE SUBFAMILY PTEROPODINAE USING PCR-RFLP ANALYSIS BASED ON PARTIAL MtDNA CYTOCHROME *b* GENE

Mohd Zairul Fairuz Bin Zakaria

Bachelor of Science with Honours
(Animal Resource Science and Management)
2006

A PHYLOGENETIC STUDY OF THE SUBFAMILY PTEROPODINAE USING PCR-RFLP ANALYSIS BASED ON PARTIAL MtDNA CYTOCHROME *b* GENE

MOHD ZAIRUL FAIRUZ BIN ZAKARIA
10548

This project is submitted in partial fulfillment of the requirements for the degree of Bachelor of Science with Honors
(Animal Resource Science and Management Program)

Faculty of Resource Science and Technology
UNIVERSITI MALAYSIA SARAWAK
2006

DECLARATION

No portion of the work referred to in this dissertation has been submitted in support of an application for another degree of qualification of this or any other university or institution of higher learning.

Mohd Zairul Fairuz Bin Zakaria

Animal Resource Science and Management Program
Faculty of Resource Science and Technology
Universiti Malaysia Sarawak

ACKNOWLEDGEMENT

In the name of Allah most Merciful and most Gracious, I would like to thank to the following people for the help with various aspects of this project: Associate Professor Dr. Mohd. Tajuddin Abdullah as my supervisor, all lecturers, Mr. Faisal Ali Anwarali Khan, Miss Ratnawati Hazali, laboratory assistants, especially Mr. Besar Ketol and Mr. Wahap Marni, postgraduate students, Andy Kho Han Guan, Fong Pooi Har, Imelda Vivian Paul, Jayaraj Vijaya Kumaran, Siti Nurlydia Sazali and course mates for their support, encouragement and contributed in this project. I wish to thank my parents Encik Zakaria Hussin and Puan Haniah Abdullah for their support and patient. I also would like wish my appreciation to Sarawak Forestry Cooperation for allowing studying bats and their facilities at Matang, Kubah and Bako national parks. My appreciation also goes to Faculty of Resource Science and Technology Universiti Malaysia Sarawak (UNIMAS) for their administrative supports.

This work is supported by IRPA research grant 09-02-09-1022-EA001 awarded to MTA and Sarawak Forestry Department provided the research permit NPW.907.4-44 to study bats in Sarawak.

Table of Contents

List of Tables

List of Figure

A PHYLOGENETIC STUDY OF THE SUBFAMILY PTEROPODINAE USING PCR-RFLP ANALYSIS BASED ON PARTIAL MITOCHONDRIAL DNA CYTOCHROME *b* GENE

Mohd Zairul Fairuz Bin Zakaria

Animal Resource Science and Management Program
Faculty of Resource Science and Technology
Universiti Malaysia Sarawak

ABSTRACT

Mitochondrial DNA variations have been extensively examined in subfamily Pteropodinae by DNA sequence analysis. The primary goal of this study was to develop genetic markers for representatives of Pteropodinae using PCR-RFLP method. This study examines 22 samples of bats from the family Pteropodidae including of *Rousettus, Pteropus* and *Cynopterus.* Total genomic DNA from different species of the Pteropodinae was extracted using CTAB method, amplified by PCR and subjected to PCR-RFLP analysis. Ten restriction enzymes used in this study are *Rsa*I, *Csp*6I, *Hae*III, *Alu*I, *Eco*RI, *Bam*HI, *Hpa*II, *Pst*I, *Sal*I and *Kpn*I. Mitochondrial DNA cytochrome *b* gene was targeted to amplify 450 bp fragments and phylogenetic relationships were reconstructed using pairwise distance of UPGMA. In general, this study is indicated that the comparison between DNA sequencing by previous study with PCR-RFLP have some advantages and disadvantages in detection of small genetic variation.

Key words: Pteropodinae, Phylogenetic, PCR-RFLP analysis, mtDNA cytochrome *b*, Restriction enzyme and UPGMA

ABSTRAK

Variasi gen mtDNA sampel kelawar dari subkeluarga Pteropodinae telah di kaji secara meluas menggunakan turutan DNA. Matlamat utama projek ini dijalankan adalah untuk mencari penanda genetic antara spesis subkeluarga Pteropodinae mengunakan kaedah PCR-RFLP. Sebanyak 22 sampel kelawar dari keluarga Pteropodidae telah digunakan termasuklah Rousettus, Pteropus, Cynopterus dan Eonycteris. Keseluruhan genom DNA dalam subkeluarga Pteropodinae diekstrak mengunakan CTAB, kemudian direplikasi menggunakan PCR dan seterusnya analisis PCR-RFLP. Sepuluh enzim restriksi yang digunakan dalam kajian ini adalah seperti RsaI, Csp6I, HaeIII, AluI, EcoRI, BamHI, HpaII, PstI, SalI dan KpnI. Sitokrom b daripada gen mtDNA disasarkan dalam lingkungan 450 bp dan hubungan filogenetik dibina menggunakan UPGMA. Secara amnya, kajian ini telah menunjukkan antara analisi turutan DNA dan PCR-RFLP mempunyai kelebihan dan kekurangnya tersendiri dalam pengesanan variasi genetik.

Kata kunci: Pteropodinae, filogenetic, PCR-RFLP analisis, mtDNA sitokrom b, enzim restriksi dan UPGMA

1.0 Introduction

Bats are known to have high capacity for dispersal and show phylogeographic patterns similar to birds (Koopman, 1994), and being the only the mammalian group with the ability to fly. Bats belong to the order Chiroptera, and include almost one quarter of all known extant mammalian species of about 916 species (Wilson and Reeder, 1993). Bats are widely distributed, and are second most numerous groups after rodents (Payne *et al.*, 1985). Approximately 88% of the species of bats are exclusively tropical and 98 species from eight families exist in Borneo Island, which make up 40% of the mammals on this island (Abdullah and Hall, 1997; Lim, 1998).

The order Chiroptera can be divided into two distinct suborders, namely, the Megachiroptera and the Microchiroptera (Findley, 1993). According to Corbet and Hill (1992), under the suborder of Chiroptera family Pteropodidae, there are 41 modern genera including 163 species distributed through out the tropics and sub tropics, from Africa, Asia, Australia region and the island of west Pacific Oceans. Pteropodids consist of all flying foxes and Old World fruit bats which are further divided into four subfamilies, namely, the diverse subfamily Pteropodinae, subfamily Macroglossinae, the monotypic subfamily Harpyionycterinae and subfamily Nyctimeninae (Corbet and Hill, 1992). Payne *et al.* (1985) stated that there are 11 genera of fruit bats currently recognized from Borneo and they fall into two subfamilies which are the Macroglossinae (nectar feeding bat) bats and the cynopterine (fruit bat). According to Andersen (1912), family Pteropodidae was divided into four subfamilies. The subfamily Pteropodinae comprises majority of the genera which includes the Epomophori that consists of *Epomophorus, Epomops* and *Nanonycteris;* Rousetti consists of *Rousettus, Pteropus* and

Dobsonia; and *Cynopterus* consists of *Myonycteris, Sphaerias, Cynopterus, Nyctimene* and *Balionycteris*. A total of 20 genera are listed in the subfamily Pteropodinae (Corbet and Hill, 1992) and in this study there are three genera of subfamily Pteropodinae that were examined, namely *Rousettus, Pteropus* and *Cynopterus* and one genus from subfamily Macroglossinae namely, *Eonycteris*. Among all the genera of Old World fruit bat, *Cynopterus* is the most diverse and broadly distributed (Corbet and Hill, 1992).

Pteropodids are relatively small to very large bats with the forearm length from 40 to 220 mm. They have large eyes and are dependent on sight and smell to navigate and find food (Mickleburgh *et al*., 1992). The members of the subfamily Pteropodinae roost in rock shelters and caves, usually near cave entrances. They usually feed on fruits and flowers and become active during dusk and dawn to search for food. Most of them do not have an echolocation system except for *Rousettus* which clicks its tongue to orient its flight path in deep and dark regions of caves (Corbet and Hill, 1992; Mickleburgh *et al.,* 1992; Bates and Harrison, 1997). Some of the characteristic of this subfamily are they have a simple external ear with the edge forming an unbroken ring and do not have a noseleaf or tragus. According to Corbet and Hill (1992), the tail membrane is usually narrow and the tail is usually short or absent. Most of them have large and flat grinding teeth (Mickleburgh *et al.,* 1992), and the muzzle and jaws are strongly built (Bates and Harrison, 1997).

This can explain why fruit bats lack echolocation and they rely instead on eyesight and smell to navigate (Payne *et al*., 1985). Fruit bats mainly feed on fruits, so their teeth are simple and rather degenerate (Payne *et al*., 1985).

1.1 Mitochondrial DNA

Several molecular methods such as gene sequencing, genomic restriction analysis and conventional ribotyping are currently available for application on differential of species (Sibley and Ahlquist, 1995). The aim of this study was to develop the genetic marker of the subfamily Pteropodinae using PCR-RFLP of cytochrome *b* (cyt *b)* from the mitochondrial DNA (mtDNA) gene (Figure 1). The mtDNA cyt *b* gene is widely used in systematic studies to resolve the divergence at many taxonomic levels (Irwin *et al.,* 1991). It is one of the most useful genes for phylogenetic work (Brown *et al.,* 1982).

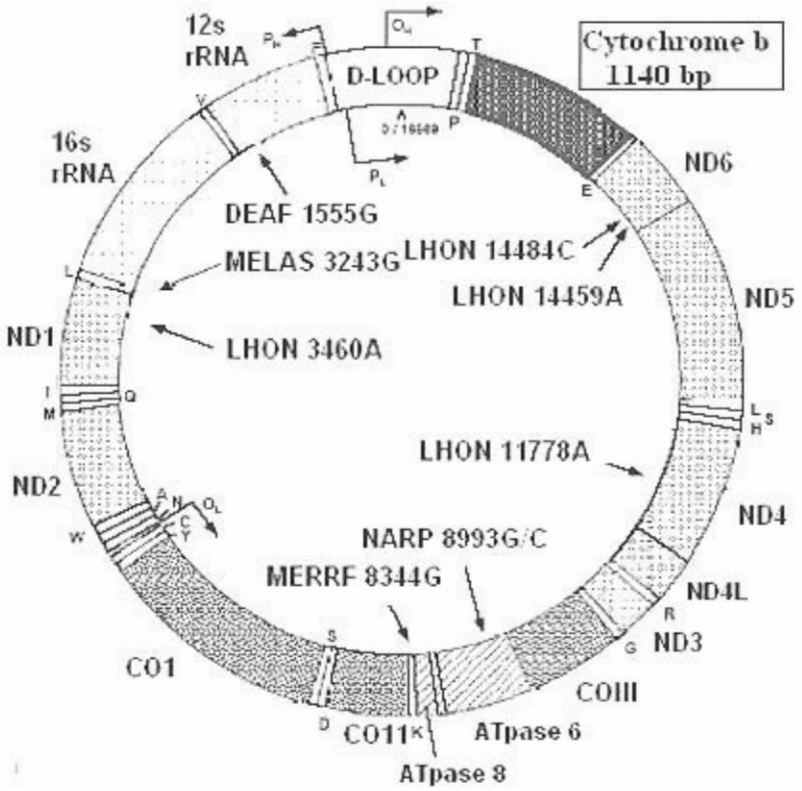

Figure 1: The location of cyt *b* in mtDNA (Source: Anon, 2004)

The mtDNA cyt *b* gene was chosen as a region because of its utility for investigating phylogenetic relationships at varied hierarchical level and playing an important role as a genetic marker (Palumbi, 1996). According to Garrett and Grisham (2002), mtDNA encode gene products required for mitochondrial protein synthesis, electron transport and oxidative phosphorylation. The cyt *b* gene was known important in the respiratory chain of cellular metabolism (Irwin *et al.*, 1991). According to Harrison (1989), simple sequence organization, maternal inheritance and absence of recombination make mtDNA an ideal marker for maternal genealogies and it evolves faster than nuclear DNA.

In recent years, the homologous segments of mtDNA was successfully amplified from the polymerase chain reaction (PCR) using a universal cyt *b*. PCR is an enzymatic synthesis and it will amplify of DNA between 50 bp and 5000 bp in size by flanking regions of target DNA sequences (Scott and Graham, 2001). The amplification of a DNA fragment is flanked with two oligonucleotide primers hybridizing to opposite strands of the target and it enables small amounts of specific DNA fragment to be amplified more than a million-fold (Scott and Graham, 2001).

1.2 Molecular analysis

Recently, different DNA region marker systems have been proposed to be used in molecular studies as an additional tool to detect the genetic diversity. The three most common types of markers used today are Restriction Fragment Length Polymorphism (RFLP), Random Amplified Polymorphism Development (RAPD) and isozymes (Lovette *et al.*, 1999). Of the three marker types, RFLP have been used the most extensively. RFLP is a technique in which

organisms may be differentiated by analysis of patterns derived from cleavage of their DNA using restriction endonuclease (Palumbi, 1996). Restriction endonucleases are enzymes that cleave DNA molecules at specific nucleotide sequences depending on the particular enzyme used and recognition sites are usually four to six base pairs (bp) in length (Scott and Graham, 2001). The restriction enzymes are useful for detecting genetic variation and double stranded cut at specific base sequences (Russel, 1997).

The study conducted by Ferreira *et al.* (2005) found that two closely related of bat species, namely *Platyrrhinus lineatus* and *P. recifinus* could be discriminate by using PCR-RFLP. A major use of molecular markers currently is for constructing genetic maps by analysing the co-segregation of markers and phenotypes or traits in defined populations (Russel, 1997). There are several advantages of using PCR-RFLP in comparison with the RAPD and isozyme markers: 1) they are codominant and unaffected by the environment; 2) any source DNA can be used for the analysis; and 3) many markers can be mapped in a population that is not stressed by the effects of phenotypic mutations (Springer *et al.*, 1992).

1.3 Problem statement

Bats have several characteristics that make them vulnerable to disturbance. Many species roost in large colonies, an adaptation that renders large numbers of individual bats vulnerable to the same catastrophe (Pierson and Raine, 1992). For some species, there are only a few roosts for all the individuals of one species, making them particularly vulnerable (Pierson and Raine, 1992). Bats are important for pollination and are agents of seed dispersal in tropical forest such as of durians (Mickleburgh *et al.*, 1992). Francis (1994) found that the diversity of

fruit bats in the forest canopy is declining. The factors of declinations of bats population in recent time are because of habitat loss, heavily hunting and the logging activity (Mickleburgh *et al.*, 1992).

Some of species bats are also difficult to identify by using available works of reference and sometimes lead to misidentification among closely related species. According to Campbell *et al.* (2004), there were three nominal closely related species in the genus *Cynopterus*: *Cynopterus horsfieldi*, *C. sphinx* and the *C. brachyotis*. Examinations of teeth of the bats are crucial to examined to confirm their identities in addition to coloration, measurement of forearm and size of ear (Payne *et al.*, 1995). However, some confusion between immature and adult specimen with similar morphology also contribute towards wrong identification. Closely related species can be identified through DNA sequencing technique and the results are more significantly accurate (Robinson *et al.*, 2001). Abdullah (2003) and Campbell *et. al* (2004), observed genetic divergence within taxa *C. brachyotis* populations by using mtDNA sequences.

The DNA sequencing is expensive compared to RFLP. The molecular study of bats especially using PCR-RFLP technique is also lack in identification the closely related species in Malaysia.

2.0 Objectives

The main objective of conducting this study is to identify genetic marker and analysis the phylogenetic relationship between species in the subfamily Pteropodinae. The secondary objective to make comparison with the data obtained previously using DNA sequencing method with current study using PCR-RFLP technique.

3.0 Materials and Methods

3.1 Samples Collection

A total of 22 bats tissue samples were collected from Bako National Park, Kubah National Park, Matang Wildlife Centre and additional specimen of bats was obtained from the Zoological Museum Universiti Malaysia Sarawak (UNIMAS) (Table 1). Bat specimens were preserved in 75% of ethanol. The tissue samples were excised from the pectoral muscle of bats.

Table 1: Species of bats and their localities used in this study

No	Species	Reference	Locality (site, state)
1	*R. amplexicaudatus*	Serasot 1	Serasot, Sarawak
2	*Pteropus vampyrus*	SR2	Serian, Sarawak
3	*P. vampyrus*	PP1	Pulau Patok Patok, Sarawak
4	*C. brachyotis*	BNP 241	Bako National Park, Sarawak
5	*C. brachyotis*	BNP 229	Bako National Park, Sarawak
6	*C. brachyotis*	SD40	Sungai Dusun, Selangor
7	*C. brachyotis*	JC75	Jambusan Cave, Sarawak
8	*C. brachyotis*	WK2	Wang Kelian, Perlis
9	*C. brachyotis*	TL8	Wang Kelian, Perlis
10	*C. sphinx*	71	Taiping, Perak
11	*C. sphinx*	72	Wang Kelian, Perlis
12	*C. sphinx*	75	Bukit Jernih, Perlis
13	*C. sphinx*	77	Selama, Perak
14	*C. horsfieldi*	SD25	Sungai Dusun, Selangor
15	*C. horsfieldi*	SD24	Sungai Dusun, Selangor
16	*C. horsfieldi*	TL6	Wang Kelian, Perlis
17	*C. horsfieldi*	MTA 96318	Jambusan Cave, Sarawak
18	*C. horsfieldi*	MTA96323	Kubah National Park, Sarawak
19	*Eonycteris spelaea*	BNP 014	Bako National Park, Sarawak
20	*E. spelaea*	BNP 009	Bako National Park, Sarawak
21	*E. spelaea*	UNIMAS510	Tawau Hills National Park, Sabah
22	*E. major*	UNIMAS130	Gunung Pueh, Sarawak

3.2 DNA Extraction and Isolation

Total genomic DNA of fruit bats was extracted from muscle tissue using modified C-TAB (cetyl-tri-methylammonium bromide) protocol following Grewe *et al.* (1993). Tissue samples, were diced and resuspended in 700 µl of 2X CTAB with 5 µl Proteinase-K (Scott and Graham, 2001).

Microcentrifuge tubes containing tissue samples and CTAB were incubated in the waterbath at 60°C for one hour or longer until the entire tissue was completely dissolved. After incubation, 700 µl of isoamyl-chloroform alcohol (24:1) was added (Grewe *et al.*, 1993). Then, the microcentrifuge tubes were centrifuged for 10 minutes at 13,000 rpm to separate the layer between DNA and the other cellular components. This process until the interphase between upper aqueous and organic phase was cleared. Upper part of aqueous phase was pipette out about 500 µl and transferred into a new microcentrifuge tube containing 600 µl 100% ethanol (EtOH). The microcentrifuge tube was centrifuged for 10 minutes (Grewe *et al.*, 1993). The pallet of DNA washed with 600µl cold 70% ethanol and 25µl 3M NaOH and centrifuge again at 13000 rpm for 10 minute. The ethanol was discarded and the pellets left in the microcentrifuge tubes until it become air dried in the air. After a few minutes, the dried pellets were resuspended in de-ionized water about 40 µl or depending on the size of the pellet and stored at -20°C until further use.

The presence of DNA extraction was determined by electrophoresis of 1% agarose gel containing ethidium bromide and 1kb DNA ladder run at 90V for 45 minutes. The result was visualized by using UV transilluminator and photograph of DNA bands was taken on Polaroid

film. According to Scott and Graham (2001), the using of agarose gel electrophoresis is one technique to separate and identify protein molecules that differ by as little as a single amino acid.

3.3 PCR Amplification

Mitochondrial DNA (mtDNA) amplification was performed using cyt *b* primer combination of GLUD G-L as forward primer and CB2-H as reverse primer (Table 2). Approximately 450 bp fragments of the mtDNA cyt *b* gene were amplified using standard PCR procedures using Biometra T-Personal thermocycler machine.

Table 2: Sequence for two partial cytochrome *b* primers (Palumbi, 1996)

Primer	Sequence	Direction
GLUD G-L	5'-TGA CTT GAA RAA CCA YCG TTG - 3'	Forward
CB2-H	5'-CCC TCA GAA TGA TAT TTG TCC TCA -3'	Reverse

Double stranded DNA amplification were performed in 25 µl volume, containing 16.8 µl dH_2O, 2.5 µl reaction buffers, 0.5 µl dNTP, 1.25 µl for each forward and reverse primer, 1.0 µl template of DNA and 0.2 µl *Taq* polymerase. The master mix (Table 3) was prepared in 1.5 ml microcentrifuge tube and mixed well by reversion to bring all reaction components to the bottom.

Table 3: PCR ingredient mixture

Component	1 reaction
dH2O	16.8 µl
10 x reaction buffer	2.5 µl
dNTP mix	0.5 µl
Primer (Glud G-L)	1.25 µl
Primer (CB2-H)	1.25 µl
Template DNA	1.0 µl
Taq polymerase	0.2 µl
$MgCl_2$	1.5 µl
Final volume	**25.0 µl**

DNA amplification was performed in a programmable thermal cycler (Biometra T-personal) using the following profile as shown in Table 4.

Table 4: PCR profile and temperature

Step	Temperature (°C)	Time (minutes)	Cycles
1. Initial denaturation	94^0C	2	1
2. Strand denaturation	94^0C	1	
3. Annealing	56^0C	1	30
4. Primer extension	72^0C	2	
5. Final extension	72^0C	5	1
6. Soaking	4^0C	∞	

3.4 Restriction Fragment Length Polymorphism (RFLP)

The ingredient of PCR-RFLP mastermix (Table 5) for 10 µl reaction mixtures containing 1.0 µl 10x buffer (Fermentas), 0.5 µl of restriction enzyme, 4.5 µl dH2O and 4.0 µl of PCR products. The samples were incubated at 37°C in waterbath for three or four hours. RFLP products were run on 2.0% agarose gel electrophoresis for three hours at 70V containing

ethidium bromide. The RFLP fragments were visualized under the UV-transilluminator and the photograph of the gel was taken using Polaroid film.

Table 5: PCR-RFLP reaction mixture

Cocktail	**1 reaction**
10 x buffer	1.0 µl
Restriction enzyme	0.5 µl
dH_2O	4.5 µl
PCR product	4.0 µl
Final volume	**10.0 µl**

The PCR products segments from each specimen were subsequently screened for polymorphism with the following 10 restriction endonucleases (Table 6).

Table 6: Restriction enzyme and the cleavage sites used in this study

Enzyme	**Cleavage sites**
*Rsa*I	GT↑AC
*Csp*6I	G↑TAC
*Hae*III	GG↑CC
*Alu*I	AG↑CT
*Hpa*II	C↑CGG
*Eco*RI	G↑AATTC
*Bam*HI	G↑GATCC
*Pst*I	CTGCA↑G
*Sal*I	G↑TCGAC
*Kpn*I	GGTAC↑C

3.5 Phylogenetic analysis

Restriction site data matrix was generated in which presence of a restriction site is coded as 1 and the absence as 0. From this data matrix, the Nei and Li coefficient (Figure 2) of genetic similarity (Nei and Li, 1979) was calculated between each pair of genotypes after the

optimality criterion was set the distance. Data analysis was performed using the Molecular Evolutionary Genetics Analysis (MEGA) version 3.0 and unweighted pair-group method with arithmetic mean (UPGMA) (Kumar *et. al.*, 2001) to reconstruct a dendogram phylogenetic relationship of Pteropodinae.

$$F = \frac{2\,Mab}{(Ma+Mb)}$$

F = Pairwise distance
Mab = the number of fragments shared by two samples (a and b)
Ma = total number of fragments resolved in samples 'a'
Mb = the number resolved in 'b'

Figure 2: The formula of Nei and Li, 1979 (Source: Anon, 2006)

4.0 Results

4.1 Sample extraction

DNA extraction from the fresh samples produced a much higher quality of DNA, based on appearance of a single and bright band on Polaroid film. Compared to DNA extraction of old samples, fresh samples show a bright band. The quality of the isolated DNA was observed from the appearance of a single and bright band above the position of 1 kb ladder (Figure 3).

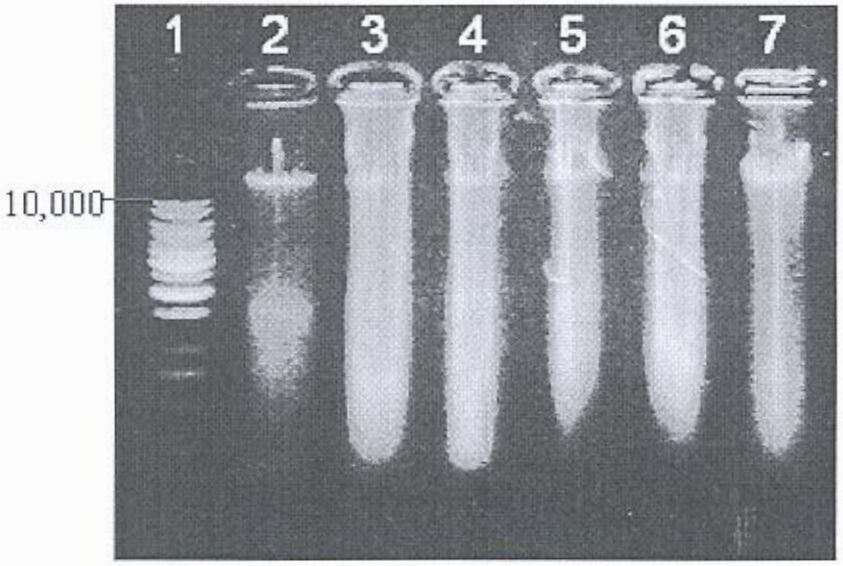

Figure 3: DNA extraction. Lane 1 represent GeneRuler 1kb DNA ladder as a standard size marker. Line 1= 1 kb DNA Ladder; Lane 2= DNA from *E. spelaea* (BNP 014); Lane 3= DNA from *E. spelaea* (BNP 009); Lane 4= DNA from *C. brachyotis* (BNP 241); Lane 5= DNA from *C. brachyotis* (BNP 229); Lane 6= DNA from *C. brachyotis* (SD40); Lane 7= DNA from *C. brachyotis* (TL8).

The samples which obtained from the UNIMAS Museum showed smearing bands because the DNA might have been degraded and cause by the present of RNA or the samples was not well preserve (Figure 4).

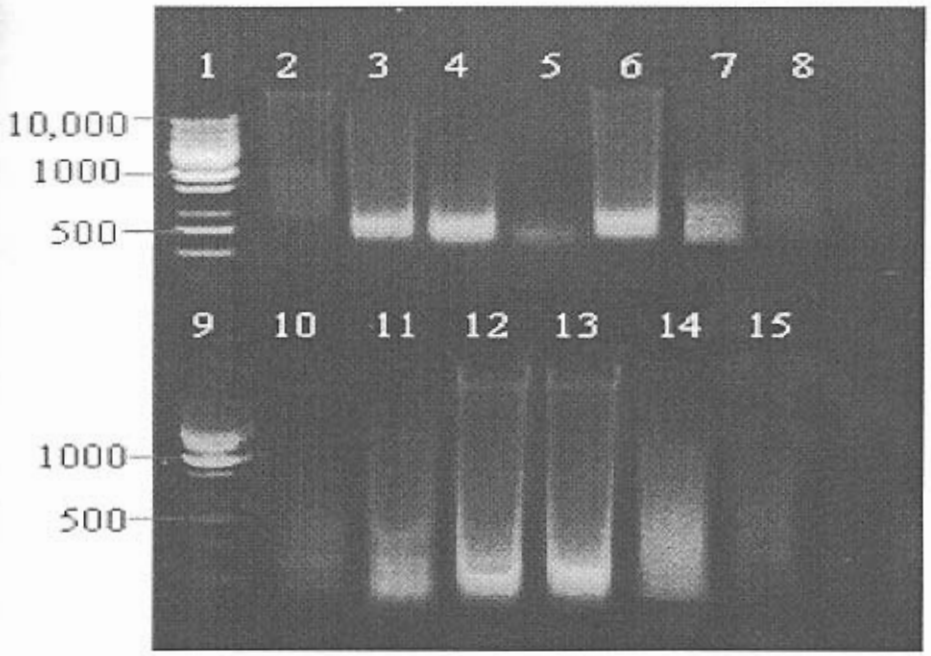

Figure 4: Lane 1 and 9 represents 1kb DNA Ladder, while lane 2 to 15 represents the extraction products for the old samples whereas some of them show smearing result. Lane 1 and 9= 1 kb DNA Ladder; Lane 2= DNA from *C. horsfieldi* (MTA 96318); Lane 3= DNA from *C. horsfieldi* (MTA 96323); Lane 4= DNA from *C. horsfieldi* (TL 6); Lane 5= DNA from *C. sphinx* (75); Lane 6= DNA from *C. sphinx* (71); Lane 8= DNA from *C. sphinx* (72); Lane 10= DNA from *P. vampyrus* (SR2); Lane 11= DNA from *P. vampyrus* (PP1); Lane 12= DNA from *R. amplexicaudatus* (Serasot1); Lane 13= DNA from *E. spelaea* (UNIMAS 510); Lane 14= DNA from *E. major* (UNIMAS 130); Lane 15= DNA from *C. sphinx* (77).

A few tissue samples of *C. horsfieldi*, *C. sphinx* and *P. vampyrus* did not completely dissolve after one hour of incubation at 60°C due to high amount of tissue used. Longer time was needed to ensure the remaining tissue was dissolved completely. A single and bright band should appear above the range of the 1 kb DNA ladder.